APPLICATIONS OF BAYESIAN NETWORKS IN ECOLOGICAL MODELLING

Reggie Mead, John Paxton, Rick Sojda
Montana State University - Bozeman
Computer Science Department, Northern Rocky Mountain Science Center
Bozeman, MT 59717 USA
mead@cs.montana.edu, paxton@cs.montana.edu, sojda@montana.edu

ABSTRACT
Bayesian belief networks are a popular tool for reasoning under uncertainty. Certain advantages make them well suited for applications in ecological modelling. In this paper, we provide an overview of Bayesian belief networks and offer examples of their use in ecological modelling. We also review hierarchical Bayesian modelling and influence diagrams.

KEY WORDS
Bayesian Belief Networks, Modelling and Simulation of Ecosystems, Statistics

1. Introduction

Ecological modelling often involves working with complex systems operating under uncertain conditions. Over the past half century, Bayesian methods have emerged as a preferred method for reasoning with uncertainty due to their mathematical foundation. Although Bayesian theory does not solve all problems in probabilistic reasoning, it has given scientists a sound framework within which uncertainty can be represented and analyzed pragmatically. By looking at systems probabilistically, the models constructed explicitly represent the uncertainty in the underlying system.

1.1 Bayesian Methodology

The Bayesian methodology is built upon the well known Bayes' Rule, which is itself derived from the fundamental rule for probability calculus.

$$P(a,b) = P(a \mid b) * P(b) \qquad (1)$$

In Equation 1, P(a,b) is the joint probability of both events a and b occurring, P(a|b) is the conditional probability of event a occurring given that event b occurred, and P(b) is the probability of event b occurring.

Although not included here, further derivation produces Bayes' rule [1].

$$P(b \mid a) = \frac{P(a \mid b) * P(b)}{P(a)} \qquad (2)$$

Bayes' rule not only opens the door to systems that evolve probabilities as new evidence is acquired, but also, as will be seen in the next section, provides the underpinning for the inferential mechanisms used in Bayesian belief networks [1].

Despite its benefits, the Bayesian approach also has drawbacks. One drawback is the difficulty of obtaining accurate conditional probabilities. When adequate data is unavailable, sometimes experts must estimate the missing probabilities subjectively [2]. Another drawback is that the approach can be computationally intensive, especially when the variables being studied are not conditionally independent of one another.

1.2 Bayesian Belief Networks

A Bayesian belief network (BBN) [1] is a directed acyclic graph (DAG) that provides a compact representation or factorization of the joint probability distribution for a group of variables. Graphically, a BBN contains nodes and directed edges between those nodes. A simple illustration is provided in Figure 1. Each node is a variable that can be in one of a finite number of states. The links or arrows between the nodes represent causal relationships between those nodes. All of the variables in Figure 1 are Boolean variables, but there is no restriction on the number of states that a variable can have. Because the absence of an edge between two nodes implies conditional independence, the probability distribution of a node can be determined by considering the distributions of its parents. In this way, the joint probability distribution for the entire network can be specified. This relationship can be captured mathematically using the chain rule in Equation 3 [3].

$$p(x) = \prod_{i=1}^{n} p(x_i \mid parents(x_i)) \qquad (3)$$

In general terms, this equation states that the joint probability distribution for node x is equal to the product of the probability of each component x_i of x given the parents of x_i. Each node has an associated conditional probability table that provides the probability of it being in a particular state, given any combination of parent states. When evidence is entered for a node in the network, the fundamental rule for probability calculus and Bayes' rule can be used to propagate this evidence through the network, updating affected probability distributions. Evidence can be propagated from parents to children as well as from children to parents, making this method very effective for both prediction and diagnosis [1, 3].

ecology and natural resources. Several other examples are listed in Table 1 at the end of this section.

2.1 A BBN for Eutrophication Modelling

One example of how a BBN might be used in ecological modelling is given by Borsuk et al. [5]. In this paper, a BBN is used in an eutrophication model. The network produced was capable of synthesis, prediction, and uncertainty analysis.

Scientists were interested in understanding the system of eutrophication that was taking place in the Neuse River estuary in North Carolina. Decision makers were considering new legislation concerning the total maximum daily load for nitrogen, a known major cause of eutrophication. They were therefore interested in quantifying the relationship between nitrogen loading and variables of interest, including
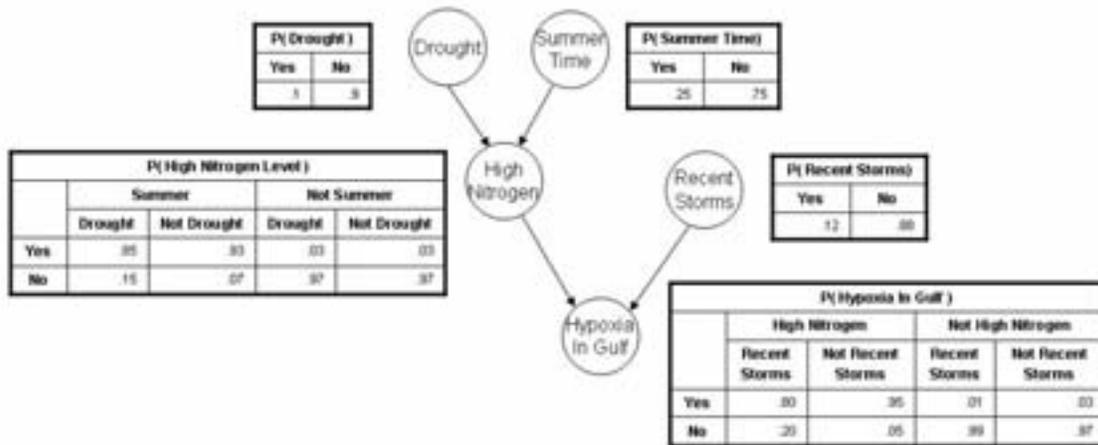

Figure 1 A BBN Modelling Hypoxia

The biggest problem with using a BBN is that exact or even approximate inference in an arbitrary network is NP-Hard in time complexity [4]. In other words, there is no known polynomial time algorithm that can provide the inference. Instead, exact inference requires time that is exponential in the number of variables. Networks with more than just a few nodes quickly become intractable to use.

2. Ecological Examples

The following two examples illustrate the use of BBNs in ecological modelling. BBNs are versatile and have been used to facilitate many different forms of probabilistic reasoning in

shellfish population size, size and frequency of algal blooms, size and frequency of fish kill, and others. The available knowledge related to this problem existed in a number of different forms. It included knowledge from process sub-models, knowledge from regression sub-models, and general knowledge held by experts. Likewise, the knowledge also existed at a variety of different scales. A BBN was used to integrate these sub-models and disparate knowledge.

To develop the network, a comprehensive survey of the relevant literature was performed and a number of meetings with experts were conducted to identify variables that should be represented as nodes in the BBN. After this process concluded, the authors developed a network with 35 nodes and 55 links. In an

attempt to make the network more tractable, additional analysis was performed to eliminate nodes that were irrelevant or unrelated to nitrogen. Other nodes were eliminated for being uncontrollable, unpredictable, or unobservable at an appropriate scale. This simplification reduced the number of nodes from 35 to 14 and the number of links from 55 to 17. A number of the remaining variables were described by sub-models including algal density, pfiesteria abundance, carbon production, sediment oxygen demand, bottom water oxygen concentration, shellfish survival, fish population health, and fish kills. The final model structure is illustrated in Figure 2 [5].

Rather than storing the conditional probabilities for each node in a conditional probability table, the authors used an alternative approach whereby each node has a corresponding function that produces the probability distribution for that node. This function was in the form of $X=f(p, \theta, \varepsilon)$ where p are the parents of x, θ are parameters relating p and x, and ε is an error term. This functional form allowed the p, θ, and ε terms to be specified in a variety of ways, making it possible to select the best approach on a per node basis, taking into account the amount and kind of data available for each of the submodels.

After all initial conditional probabilities were established, different scenarios for nitrogen loading were entered into the network and marginal probability distributions for variables of interest were estimated using Monte Carlo [6] or Latin Hypercube [7] sampling. Although the resulting model produced useful predictions for decision makers and the results of the model were favorable when compared with data, the authors' objective was not to produce a model that more realistically represented the actual system, but that instead more realistically represented what was known about the system. This integration of various forms of knowledge at various scales was simplified by the use of a BBN.

This study identified several drawbacks of BBNs. The most significant drawback is the inability of a BBN to adequately capture the often dynamic nature of the systems being modeled. Specifically, the requirement that BBNs are directed acyclic graphs dictates that they are incapable of representing system feedback. This limitation might lead to poor results in systems where dynamic processes like feedback play a significant role.
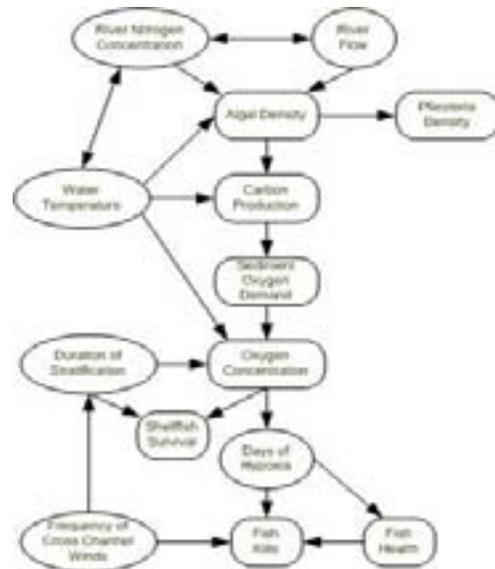

Figure 2 A BBN Modelling Eutrophication

Another drawback is that BBNs do not in themselves offer a solution to the problem of representing structural uncertainty. The uncertainty in the causal structure of the network is unaccounted for, leading to model predictions that underestimate the level of uncertainty.

2.2 A BBN for Modelling Ecological Webs

Marcot et al. [8] offers an example where BBNs are used to model the causal web between biotic factors, habitat conditions, and management for some vertebrate and invertebrate species in the Columbia River Basin. This paper follows a similar approach to that described in the previous subsection for constructing and parameterizing the model. Both current literature and expert judgment were used. One difference between the two projects is that this paper is not primarily concerned with the effect that a single controlled variable (nitrogen loading, for example) has on a few primary variables of concern (e.g. fish kills or health and shellfish abundance), but is more interested in discovering and quantifying the relationships between many of the nodes in the network that often represent key environmental correlates.

Two separate BBN groups were used. These BBNs were eventually extended into influence diagrams (section 3.2). The first was used for aquatic wildlife and the second was used for terrestrial wildlife. The extension to influence diagrams allowed optimal pathways through the network to be made explicit and helped prioritize

Authors	Title / Publication	Publication Date
P. Bacon, J. Cain & D. Howard	Belief network models of land manager decisions and land use change	2002
	Journal of Environmental Management	
M. Borsuk, P. Reichert, A. Peter, E. Schager & P. Burkhardt-Holm	Assessing the decline of brown trout (Salmo trutta) in Swiss rivers using a Bayesian probability network	2006
	Ecological Modelling	
C. Smith & O. Bosch	Integrating disparate knowledge to improve natural resource management	2004
	ISCO 2004	

Table 1 Other Examples of BBNs in Ecology

the network attributes being monitored. Sensitivity analysis was used to determine which attributes of the model had the most significance.

The two BBN model groups were developed at a variety of scales. The aquatic group was developed at two scales, the first consisting of habitat and other biotic influences and the second consisting of landscape properties and management activities. The models in the terrestrial group were developed at three different scales. The first was site-specific, the second was sub-watershed, and the third was developed at the basin scale. The resulting model was able to identify which key environment correlates had the biggest effect on local population response.

The greatest benefit of using a BBN in this study resulted from requiring experts to articulate what they knew regarding the subject. This opening of communication channels was tremendously helpful for understanding the problem being investigated. It was important that the knowledge used to construct the model be peer reviewed because personal bias can easily be built into a BBN, as it can be in other knowledge-based methods.

A cautionary note to remember is that although BBNs can combine many different forms of knowledge, it is important to remember that without any empirical data, the models provide little advantage over an educated guess. This potential to overstate expert opinion demands that BBNs be used responsibly and ethically, as is true of other knowledge-based methods.

3. Other Approaches

3.1 Hierarchical Bayesian Modelling

Parameter estimation is a common requirement when building mathematical and statistical models [9]. Typically, if parameters are identifiable, they can be accurately estimated from observation data, assuming an adequate amount of data is available. Unfortunately, this assumption is often invalid, and it is common to have sparse data for a system of interest but still be faced with the daunting task of parameterizing the model. An obvious pitfall when parameterizing a model using sparse data is the potential for overfitting the model to the data. This is always a possibility when relying on site-specific data.

An alternative to strictly using site-specific data is the exploitation of observation data for similar systems, which are often available. By combining the data from the specific system with data from similar systems, the site specific parameters become globally specific parameters. This avoids overfitting but at the cost of potentially overgeneralizing the model by assuming that parameters are shared between systems. The quest to find a compromise between site-specific and globally specific parameters led to the development of hierarchical Bayesian modelling.

Hierarchical Bayesian modelling allows each system to have its own parameters, but these parameters can be influenced by commonalities between the systems. This approach often draws on the belief that many groups of systems have possibly unique parameters for each individual system, but that these parameters are drawn from the same probability distribution. Thus, multi-system data can be used to implicitly or explicitly identify this distribution and site-specific data can be used to fine tune the parameters on a per system basis [9, 10]

Hierarchical modelling has been used with mixed results. Bayesian methods, however, have

given the approach a sound mathematical basis by using probability distributions and Bayes' rule. Cross-system data can be used to provide prior probability distributions for parameters which can then be combined with local data using Bayes' rule to produce posterior distributions.

Although hierarchical models often produce wider, less precise posterior probability distributions than global models, it is believed that in many cases this reduced precision more accurately represents the knowledge of site-specific attributes. By making this uncertainty explicit in the results, it is less likely that a user will be misled than when using a global model that assumes common parameters between systems and produces very precise but inaccurate results when these assumptions are not valid.

3.2 Influence Diagrams

Influence diagrams, an extension of Bayesian belief networks, can also be valuable in ecological modelling, especially with respect to decision making, which is often a driving force behind ecological modelling. Influence diagrams extend BBNs by adding utility nodes and decision nodes to the network. Utility nodes are used to assign value, or utility, to particular outcomes represented by a node being in a certain state. Decision nodes represent controllable decisions that have an effect on the system. Neither decision nodes nor utility nodes have a corresponding conditional probability table [1, 11].

A simple example is illustrated in Figure 3. In this diagram, ovals represent regular chance nodes, squares represent decision nodes, and rectangles with rounded edges represent utility nodes (diamonds are also common shapes for utility nodes). In this example, the trail condition, which is treated probabilistically, and the opening date, which is treated as a decision,
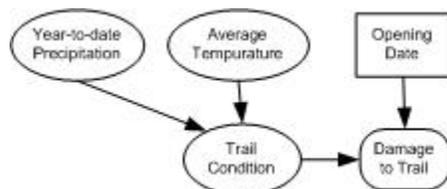

Figure 3 An Influence Diagram

both affect the amount of damage to the trail. This last property has a clear utility in terms of maintenance cost. An obvious application for this influence diagram would be determining the

opening date that results in the least damage to the trail.

The term Bayesian belief network is sometimes used interchangeably with either *influence diagram* or *graphical model* depending on the community in which it is being used. This paper takes the approach most common in the computer science community and draws a distinction between BBNs and influence diagrams, the distinction being that only the latter is allowed to have utility and decision nodes.

3.3 State of the Art

Many advances have been made that make BBNs more efficient and more effective. Most noteworthy are Markov chain Monte Carlo simulations, hierarchical and object oriented Bayesian networks, interval probability theory, and dynamic Bayesian networks.

Markov Chain Monte Carlo (MCMC) techniques are used to estimate posterior probability distributions. By using approximate inferencing, networks with more than a few nodes become tractable. MCMC techniques build a Markov chain of possible states where each state represents a unique configuration of the network. It can be shown that given enough running time, the fractional time spent in a given state is equal to the posterior probability of that state occurring [6, 12]. While MCMC techniques are not new, advances continue to be made.

Hierarchical Bayesian Networks (HBNs) [13] and Object Oriented Bayesian Networks (OOBNs) [14] are two extensions to BBNs intended to increase their ability to handle systems and processes with large and complex structures. These extensions allow the nodes of the network to themselves be instances of other networks. In this way, the causal structure can be defined on a number of different scales. OOBNs also allow classes of networks to be defined and this allows for techniques such as inheritance and encapsulation that reduce the amount of work involved in designing large networks. One advantage of using one of these extensions is the improved inferencing efficiency that results from the additional structure information.

Interval probability theory (IPT) [15, 16] can be used to express the uncertainty in the prediction itself. It does this by separating the support for a proposition from support for the negation of the proposition. In this manner, IPT supports the ability to express ambiguity in

probabilistic predictions or estimates. This can be particularly useful when eliciting expert judgments from participants that are hesitant to commit to a single probabilistic estimate. Instead, the participant is allowed to express indecision or even ignorance on a subject.

A Dynamic Bayesian Network (DBN) [17, 18] is an extension to a BBN that represents a probability model that can change with time. DBNs are also compact representations of hidden Markov models. DBNs offer a number of improvements over BBNs such as relaxing some of the feedback restrictions typical of the standard directed acyclic graphs used for BBNs. The downside to using a DBN is that the complexity tends to be greater than for static BBNs and exact inferencing is even less tractable. Instead, approximation algorithms that are often quite complex must be used.

Although some of these techniques have only recently appeared in the ecological modeling literature, their potential application to ecological systems is readily apparent.

4. Conclusions

A Bayesian belief network offers a sound mathematical framework within which probabilistic reasoning using uncertain and varying data can be performed. Its ability to combine various forms of knowledge and to evolve as new knowledge is acquired allows it to produce informed results at various levels of scale. The probabilistic nature of a BBN allows it to explicitly represent uncertainty. New computational methods and techniques keep increasing a BBN's abilities and range of practical applications.

References

[1] F. Jensen, *Bayesian networks and decision graphs* (New York: Springer-Verlag, 2001).
[2] K. Reckhow, Bayesian approaches in ecological analysis and modelling, *The role of Models in Ecosystem Science*. Princeton University Press, 2002.
[3] D. Heckerman, A tutorial on learning bayesian networks, *Microsoft Technical Report* 95-06, 1996.
[4] D. Heckerman & M. Wellman, Bayesian networks, *Communications of the ACM* 38:3 (1995): 27-30.
[5] M. Borsuk, C. Stow & K. Reckhow, A bayesian network of eutrophication models for

synthesis, prediction, and uncertainty analysis, *Ecological Modelling* 173 (2004): 219-239.
[6] A. Smith & G. Roberts, "Bayesian computation via the gibbs sampler and related markov chain monte carlo methods, *Journal of the Royal Statistical Society. Series B (Methodological)* 55:1 (1993): 3-23.
[7] J. Helton & F. Davis, Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems, *Reliability Engineering and System Safety* 81 (2003): 23-69.
[8] B. Marcot, R. Holthausen, M. Raphael, M. Rowland & M. Wisdom, Using bayesian belief networks to evaluate fish and wildlife population viability under land management alternative from an environmental impact statement, *Forest Ecology and Management* 153 (2001): 29-42.
[9] M. Borsuk, D. Hidgon, D, C. Stow & K. Reckhow, A Bayesian hierarchical model to predict benthic oxygen demand from organic matter loading in estuaries and coastal zones, *Ecological Modelling* 143 (2001): 165-181.
[10] V. Tresp, V & K. Yu, An introduction to nonparametric bayesian modelling with a focus on multi-agent learning, *Switching and Learning*. Berlin: Springer-Verlag, 2005.
[11] R. Shachter, Evaluating influence diagrams, *Operations Research* 34:6 (1986): 871-882.
[12] W. Hastings, Monte carlo sampling methods using markov chains and their applications, *Biometrika,* 57:1 (1997): 97-109.
[13] E. Gyftodimos & P. Flach, Hierarchical bayesian networks: an approach to classification and learning for structured data, *Methods and Applications of Artificial Intelligence: Third Hellenic Conference on AI, SETN 2004*, Samos, Greece, 2004.
[14] D. Koller, D. & A. Pfeffer, Object-oriented bayesian networks, *Proceedings of the Thirteenth Annual Conference on Uncertainty in Artificial Intelligence (UAI-97),* Providence, Rhode Island, 1997, 302-313.
[15] J. Hall, D. Blockley & J. Davis, Uncertain inference using interval probability theory, *International Journal of Approximate Reasoning* 19 (1998): 247-264.
[16] J. Hall, C. Twyman & A. Kay, Influence diagrams for representing uncertainty in climate-related propositions, *Climatic Change* 69 (2005): 343-365.
[17] K. Murphy, *Dynamic bayesian networks: representation, inference and learning* (Ph. D. thesis, University of California, Berkeley, 2002).
[18] S. Russell & P. Norvig, *Artificial intelligence, a modern approach* (New York: Prentice Hall, 2003).